LSI 01-203

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION FOR PATENT

ON

HOST INTERFACE BYPASS ON A FABRIC BASED ARRAY CONTROLLER

BY

RUSSELL J. HENRY	BRET WEBER
2982 PENSTEMON CIR.	2521 N. TEE TIME
WICHITA, KS 67226	WICHITA, KS 67205
CITIZEN OF USA	CITIZEN OF USA

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

"Express Mail" Mailing Label Number EL 789 439 815 US

Date of Deposit: July 27, 2001

I hereby certify that this correspondence is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 C.F.R. § 1.10 on the date indicated above and is addressed to Box Patent Application, Assistant Commissioner for Patents, Washington, D.C. 20231

BY: Penny L. Flint

Penny L. Flint

HOST INTERFACE BYPASS ON A FABRIC BASED ARRAY CONTROLLER

FIELD OF THE INVENTION

[0001] The present invention generally relates to the field of electronic storage devices, and particularly to a host interface bypass on a fabric based array controller.

BACKGROUND OF THE INVENTION

[0002] Electronic data storage is one of the most important aspects of the modern economy. Most every aspect of modern life has some element that requires the storage of electronic data. Additionally, great advances have been made in the components utilized in information handling systems, such as desktop computers, servers, information appliances, convergence devices, and the like. For example, faster processors, network connection devices, memories, and the like are constantly available which enable components and the host system itself to provide increased functionality.

[0003] However, as host and drive interface bandwidth increases, a higher burden is placed on disk array controllers to increase bandwidth. Thus, advances made in other areas of electronic storage may not be realized due to these limitations. For example, typically, disk array controller architectures use peripheral connect interface (PCI) buses as the internal interconnect between functional areas. However, such buses may encounter limitations, such a connectivity, scalability and performance.

[0004] Additionally, in order to use multiple controllers to create a large storage complex, an I/O module may have to address all the host interface chips on the associated controllers, thereby making the interface very complex.

[0005] Therefore, it would be desirable to provide a host interface bypass on a fabric based array controller.

SUMMARY OF THE INVENTION

[0006] Accordingly, the present invention is directed to a host interface bypass on a fabric based array controller. The present invention may utilize a fabric interconnect, such as Infiniband, RapidIO, and the like. This may enable features such as dynamic expansion of modules, use of external memory system, scalability of memory systems and scalability and/or connectivity of array controllers to create big storage complexes.

[0007] In a first aspect of the present invention, an apparatus includes an external electronic device suitable for performing a function, a controller and a fabric connection. The controller includes at least one internal module, the internal module providing a controller function. The fabric connection communicatively connects the external device to the controller, wherein the module of the controller is directly accessible by the external electronic device.

[0008] In a second aspect of the present invention, an apparatus includes an external electronic device suitable for performing a function, a storage array controller, and a fabric connection. The storage array controller includes at least one internal module, the internal module providing a storage array controller function. The fabric connection operable connects the external device to the storage array controller, wherein the module of the controller is accessible by the external electronic device by directly converting from an external electronic device protocol to an internal storage array controller protocol suitable for communicating with the internal module.

[0009] In a third aspect of the present invention, an apparatus includes a first external device and a second external device suitable for providing a function, a controller, and a fabric connection. The controller includes at least one internal module, the internal module providing a controller function. The fabric connection includes a fabric switch communicatively connecting the first external device, the second external device and the controller, wherein the module of the controller is directly accessible via the fabric switch

to at least one of the first external device and the second external device and the fabric connection communicatively connects the first external device to the second external device.

[0010] In a fourth aspect of the present invention, a disk array controller includes a controller and a bypass line. The controller includes at least one internal module, the internal module providing a controller function. The bypass line is directly connected with an internal interconnect of the controller through a switch, so that the at least one internal module of the controller is accessible directly by an input/output module by utilizing the bypass line.

[0011] It is to be understood that both the forgoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the invention as claimed. The accompanying drawings, which are incorporated in and constitute a part of the specification, illustrate an embodiment of the invention and together with the general description, serve to explain the principles of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] The numerous advantages of the present invention may be better understood by those skilled in the art by reference to the accompanying figures in which:

[0013] FIG. 1 is a highly diagrammatic block illustration of an exemplary embodiment of the present invention wherein a high bandwidth electronic data storage device architecture utilizes an interconnect fabric to provide connectivity;

[0014] FIG. 2A is a highly diagrammatic block illustration of an exemplary embodiment of the present invention wherein a high bandwidth electronic data storage device architecture includes a bypass line to enable direct access by an input/output module;

[0015] FIG. 2B is a flow chart illustrating an exemplary embodiment of the present invention wherein a host connect protocol is converted directly to an internal protocol;

[0016] FIG. 3 is a block diagram illustrating an embodiment of the present invention wherein a fabric is utilized to enable direct access of both internal modules of a controller and access between external devices; and

[0017] FIG. 4 is a block diagram depicting an embodiment of the present invention wherein a bypass line is utilized with multiple controllers and I/O modules.

DETAILED DESCRIPTION OF THE INVENTION

[0018] Reference will now be made in detail to the presently preferred embodiments of the invention, examples of which are illustrated in the accompanying drawings.

[0019] Referring generally now to FIGS. 1 through 4, exemplary embodiments of the present invention are shown. As host and drive interface bandwidths increase, it puts a higher burden on electronic storage device controllers, such as disk array controllers, for increased bandwidth. Previous disk array controller architectures used peripheral component interconnect (PCI) as an internal interconnect between functional areas. However, such an interconnect had limitations such as connectivity, scalability and performance. A fabric interconnect may be utilized in a electronic data storage device system, such as a disk array, to provide dynamic expansion of modules, use of external memory systems, scalability of memory systems, scalability and/or connectivity of array controllers to create larger storage complexes, and the like. Further, a fabric connect may be utilized to provide a common fabric to describe both internal and external components.

[0020] Referring now to FIG. 1, an embodiment of the present invention is shown wherein a high bandwidth electronic data storage device architecture utilizes an

interconnect fabric to provide connectivity. An electronic data storage device architecture 100 may include a plurality of electronic storage devices to provide data redundancy, such as a disk array arranged in a RAID array, and the like. To provide connectivity between components, the present invention may utilize a fabric interconnect.

[0021] A fabric interconnect of the present invention provides scalability by not limiting the number of host interfaces 102 & 104 and drive interfaces 106 & 108. Additionally, the number of modules in a controller 110 & 112 is not limited by the fabric like the physical limits PCI has on the number of connections to the bus.

[0022] To provide this increased functionality, the fabric interconnect may utilize messages, which in this instance is defined as a logical unit of work, to execute transactions. For example, message semantics may support memory operations such as remote direct memory access (RDMA) reads, writes, and channel operations such as send, receive, and the like.

[0023] Messages may be further segmented into packets. Data packets are an end-to-end fabric unit of transfer, e.g. packets are the routable unit of transfer. Hardware of the present invention may provide automatic message segmentation and re-assembly via packets. Packets include headers to identify the packet destination and the desired data. For example, packets may include a local route header (LRH), global route header (GRH), base transport header (BTH), extended transport header (ExTH), intermediate data, a message payload, an invariant CRC, a variant CRC, and the like.

[0024] A host channel adapter (HCA) 114 is generally a host concept configured for computing platforms. For example, a HCA may be configured for high-end devices, such as nodes, memory array controllers, and the like. Functions of a HCA may include implementation of a link protocol in hardware, implement software verbs, as a RDMA engine, work queue engine, memory registration and protection, and the like as

contemplated by a person of ordinary skill in the art. Thus, the HCA 114 may act to implement high-end devices in the architecture of the present invention. However, there may be implementations where the full level of functionality of an HCA is not desired, such as specialized subsystems.

[0025] Target channel adapters (TCA) 116 may be configured for specialized subsystems. TCA may be configured as simpler than HCA so that it implements what is required to minimally support fabric architecture and device-specific functionality. For example, a TCA may be configured as a work queue engine, a link protocol engine, implement transport, and the like without departing from the spirit and scope of the present invention.

[0026] To ensure packet delivery within a fabric interconnect, a switch 118 may be provided. For instance, the switch may operate as a packet-switching element within a fabric subnet. Packets may be switched, for example, based on a local identifier (LID) within the local route header (LRH) as described previously. A router, which may be thought of as a switch, may also be provided to enable packet routing between systems, as will be discussed later.

[0027] Thus, by providing a fabric interconnect, a high bandwidth electronic data storage device controller architecture may be provided. For example, database transactions may be performed by using this packetization paradigm to provide dynamic expansion of modules, use of external memory systems, scalability of memory system and scalability and/or connectivity of array controllers to create larger storage complexes.

[0028] Additionally, the fabric interconnect may be expanded to include devices external to the architecture, so that both internal and external components are accessible, as shown in FIG. 2A. For instance, large storage complex architectures may use a switched fabric to connect storage controllers to provide large storage capacities. Similarly, disk array

control architectures may use a switched fabric to connect modules within the controller. By utilizing the present invention, a fabric may be employed in both a storage complex architecture and controller architectures, thereby increasing performance.

[0029] Referring generally now to FIGS. 2A through 4, exemplary embodiments of the present invention are shown. In order to use multiple controllers to create a large storage complex, an I/O module would have to address all the host interface chips on the associated controllers, making the interface very complex. However, the present invention may provide a bypass line suitable for connecting directly with the internal interconnect of the controller through a switch. For example, as shown in the system 200 depicted in FIG. 2A, a bypass line 202 is added to the controller architecture for interfacing with an I/O module. Using the bypass line 202, an I/O module may act as a host interface chip to access controller resources directly.

[0030] Thus, as shown in FIG. 2B, by utilizing the present invention, there would not be a need to convert from a host connect protocol 252 to a controller host interface protocol 254 and then have the controller convert the controller host interface protocol 254 to an internal protocol 256 as previously required. In this way, direct mapping may be achieved from host interface protocol 252 to internal control protocol 256.

[0031] For example, referring now to FIG. 3, an embodiment 300 of the present invention is shown wherein a fabric is utilized to enable direct access of both internal modules of a controller and access between external devices. The fabric may extend through the controller as well as between controllers to provide a unified system to enable direct resource control. In this way, devices external to a controller 302, such as a host system 304, server 306, second controller 308 such as from an additional storage device provided for data redundancy, and the like as contemplated by a person of ordinary skill in the art, are connected via a fabric utilizing a switch 310. Further, the fabric extends to describe modules of the controller 302, such as module one 312, module two 314 and

module N 316. Therefore, external devices, such as the host system 304, server 306, second controller 308, may directly access the modules 312, 314 and 316 without excessive converting of protocols.

[0032] Referring now to FIG. 4, an exemplary embodiment of the present invention is shown wherein a bypass line is utilized with multiple controllers and I/O modules. By utilizing a fabric used in the bypass as the fabric used as the controller interconnect, an efficient bypass of the present invention may be provided. For example, a plurality of I/O modules, such as modules utilized to provide a connection between external devices and the modules, may be provided 402. The I/O modules 402 may communicate and interface with controllers 404, 406, 408, 410, 412 and 414, in this instance, as utilized in storage systems, such as redundant electronic data storage systems 416, 418 & 420, utilizing a fabric connect. Further, a fabric connect may be used internally through a fabric interconnect. By utilizing the same fabric scheme both internally and externally, conversion of the data between protocols, as previously required, is no longer needed, thereby providing an efficient mechanism for the transfer of data. Thus, a host may directly access modules of a controller, such as a disk array controller, without converting protocols twice within the controller.

[0033] In this way, by utilizing packets, globally unique identifiers, and the like over a fabric connect both internally and externally, a flexible, scalable and dynamic electronic data storage system may be provided. For example, multiple data storage arrays may be provided to implement a storage complex, without being limited by the previous connection restraints encountered in PCI implementations. Additionally, multiple interfaces, external memory systems and array controllers may be utilized without encountering the bandwidth bottleneck of previous interconnects. Further, a single fabric connect may be provided to define both internal modules of a device and external devices. A variety of fabric connects are contemplated by the present invention without departing from the spirit and scope thereof, such as Infiniband, RapidIO and the like.

[0034] It is believed that the host interface bypass on a fabric based array controller of the present invention and many of its attendant advantages will be understood by the forgoing description. It is also believed that it will be apparent that various changes may be made in the form, construction and arrangement of the components thereof without departing from the scope and spirit of the invention or without sacrificing all of its material advantages. The form herein before described being merely an explanatory embodiment thereof. It is the intention of the following claims to encompass and include such changes.